

CLAIMS:

1. A speech processing apparatus comprising:

means for receiving a set of signal values  
5 representative of a speech signal generated by a speech  
source as distorted by a transmission channel between the  
speech source and the receiving means;

a memory for storing data defining a predetermined  
function derived from a predetermined signal model which  
10 includes a first part having first parameters which  
models said source and a second part having second  
parameters which models said channel and said function  
being in terms of said first and second parameters and in  
terms of a set of signal values;

15 means for applying said set of received signal  
values to said stored function; and

means for processing said function with those values  
applied to obtain values of said first parameters that  
are representative of said speech generated by said  
20 speech source before it was distorted by said  
transmission channel.

2. An apparatus according to claim 1, wherein said  
function generates, for a given set of received signal  
25 values, a probability density function which defines, for

a given set of first and second parameters, the probability that the predetermined signal model has those parameter values, given that the signal model is assumed to have generated the received set of signal values and wherein said processing means comprises means for drawing samples from said probability density function and means for determining said values of said first parameters that are representative of the speech from the drawn samples.

3. An apparatus according to claim 2, further comprising means for evaluating said probability density function for the set of received signal values using one or more of said drawn samples or parameter values for different numbers of parameter values, to determine respective probabilities that the predetermined signal model has those parameter values and wherein said processing means is operable to process at least some of said drawn samples of parameter values and said evaluated probabilities to determine said values of said first parameters that are representative of the speech generated by said source before it was distorted by said transmission channel.

4. An apparatus according to claim 2, wherein said processing means is operable to determine a histogram of

said drawn samples and wherein said values of said first parameters are determined from said histogram.

5        5.    An apparatus according to claim 4, wherein said processing means is operable to determine said values of said first parameters using a weighted sum of said drawn samples, and wherein the weighting is determined from said histogram.

10       6.    An apparatus according to claim 2, wherein said sampling means is operable to draw samples iteratively from said probability density function.

15       7.    An apparatus according to claim 2, wherein said sampling means is operable to draw samples of both said first and second parameters.

20       8.    An apparatus according to claim 2, wherein said sampling means comprises a Gibbs sampler.

25       9.    An apparatus according to claim 2, wherein said processing means comprises means for analysing at least some of said drawn samples of parameter values to determine a measure of the variance of said samples and the apparatus further comprises means for outputting a

signal indicative of the quality of the received set of signal values in dependence upon said determined variance measure.

5 10. An apparatus according to claim 9, wherein said probability density function is in terms of said variance measure and wherein said analysing means is operable to draw samples of said variance measure from said probability density function.

10 11. An apparatus according to claim 1, wherein said function is in terms of a set of raw speech signal values representative of speech generated by said source before being distorted by said transmission channel, 15 wherein the apparatus further comprises second processing means for processing the received set of signal values with initial estimates of said first and second parameters, to generate an estimate of the raw speech signal values corresponding to the received set of 20 signal values and wherein said applying means is operable to apply said estimated set of raw speech signal values to said function in addition to said set of received signal values.

25 12. An apparatus according to claim 11, wherein said

second processing means comprises a simulation smoother.

13. An apparatus according to claim 11, wherein said second processing means comprises a Kalman filter.

5

14. An apparatus according to claim 1, wherein said receiving means is operable to receive a sequence of sets of signal values representative of a speech signal generated by a speech source as distorted by said transmission channel and wherein said processing means is operable to obtain values of said first parameters for the speech within each set of signal values in said sequence.

10

15. An apparatus according to claim 14, wherein said processing means is operable to use the values of said first parameters obtained during the processing of a preceding set of signal values as initial estimates for the values of said first parameters for a current set of signal values being processed.

20

16. An apparatus according to claim 14, wherein said sets of signal values in said sequence are non-overlapping.

25

17. An apparatus according to claim 14, wherein said processing means comprises means for varying the number of parameter values used to represent the speech signal within each set of signal values.

5

18. An apparatus according to claim 1, wherein said first part is an auto-regressive model and said first parameters comprise auto-regressive model coefficients.

10

19. An apparatus according to claim 1, wherein said second part is a moving average model and said second parameters comprise moving average model coefficients.

15

20. An apparatus according to claim 1, further comprising means for comparing said determined parameter values with pre-stored parameter values to generate a comparison result.

20

21. An apparatus according to claim 1, further comprising recognition means for comparing said determined parameter values with pre-stored reference models to generate a recognition result.

25

22. An apparatus according to claim 1, further

comprising speaker verification means for comparing said determined parameter values with pre-stored speaker models to generate a verification result.

5        23. A speech processing method comprising the steps of:  
         receiving a set of signal values representative of  
         a speech signal generated by a speech source as modified  
         by a transmission channel between the speech source and  
         the receiver;

10        storing data defining a predetermined function  
         derived from a predetermined signal model which includes  
         a first part having first parameters which models said  
         source and a second part having second parameters which  
         models said channel, said function being in terms of said  
15        first and second parameters and generating, for a given  
         set of signal values, a probability density function  
         which defines, for a given set of first parameters and  
         second parameters, the probability that the predetermined  
         signal model has those parameter values, given that the  
20        signal model is assumed to have generated the received  
         set of signal values;

         applying said set of received signal values to said  
         function;

         processing said function with those values applied  
25        to derive samples of at least said first parameters from

said probability density function; and

means for analysing at least some of said derived samples to determine values of said first parameters that are representative of said speech signal generated by said source before it was modified by said transmission channel.

24. A method according to claim 23, wherein said function generates, for a given set of received signal values, a probability density function which defines, for a given set of first and second parameters, the probability that the predetermined signal model has those parameter values, given that the signal model is assumed to have generated the received set of signal values and wherein said processing step comprises the steps of drawing samples from said probability density function and determining said values of said first parameters that are representative of the speech from the drawn samples.

25. A method according to claim 24, further comprising the step of evaluating said probability density function for the set of received signal values using one or more of said drawn samples or parameter values for different numbers of parameter values, to determine respective probabilities that the predetermined signal model has



those parameter values and wherein said processing step processes at least some of said drawn samples of parameter values and said evaluated probabilities to determine said values of said first parameters that are representative of the speech generated by said source before it was distorted by said transmission channel.

26. A method according to claim 24, wherein said processing step determines a histogram of said drawn samples and wherein said values of said first parameters are determined from said histogram.

27. A method according to claim 26, wherein said processing step determines said values of said first parameters using a weighted sum of said drawn samples, and wherein the weighting is determined from said histogram.

28. A method according to claim 29, wherein said sampling step draws samples iteratively from said probability density function.

29. A method according to claim 24, wherein said sampling step draws samples of both said first and second parameters.

30. A method according to claim 24, wherein said sampling step uses a Gibbs sampling technique.

5 31. A method according to claim 24, wherein said processing step comprises the step of analysing at least some of said drawn samples of parameter values to determine a measure of the variance of said samples and wherein the method further comprises the step of outputting a signal indicative of the quality of the  
10 received set of signal values in dependence upon said determined variance measure.

15 32. A method according to claim 31, wherein said probability density function is in terms of said variance measure and wherein said analysing step draws samples of said variance measure from said probability density function.

20 33. A method according to claim 23, wherein said function is in terms of a set of raw speech signal values representative of speech generated by said source before being distorted by said transmission channel, wherein the method further comprises a second processing step of processing the received set of signal values with  
25 initial estimates of said first and second parameters, to

generate an estimate of the raw speech signal values corresponding to the received set of signal values and wherein said applying step applies said estimated set of raw speech signal values to said function in addition to said set of received signal values.

34. A method according to claim 33, wherein said second processing step uses a simulation smoother.

35. A method according to claim 23, wherein said second processing step uses a Kalman filter.

36. A method according to claim 23, wherein said receiving step receives a sequence of sets of signal values representative of a speech signal generated by a speech source as distorted by said transmission channel and wherein said processing step obtains values of said first parameters for the speech within each set of signal values in said sequence.

37. A method according to claim 36, wherein said processing step uses the values of said first parameters obtained during the processing of a preceding set of signal values as initial estimates for the values of said first parameters for a current set of signal values being

processed.

38. A method according to claim 36, wherein said sets of signal values in said sequence are non-overlapping.

5

39. A method according to claim 36, wherein said processing step comprises the step of varying the number of parameter values used to represent the speech signal within each set of signal values.

10

40. A method according to claim 23, wherein said first part is an auto aggressive model and said first parameters comprise auto aggressive model coefficients.

15

41. A method according to claim 23, wherein said second part is a moving average model and said second parameters comprise moving average model coefficients.

20

42. A method according to claim 23, further comprising the step of comparing said determined parameter values with pre-stored parameter values to generate a comparison result.

25

43. A method according to claim 23, further comprising

the step of comparing said determined parameter values with pre-stored reference models to generate a recognition result.

5        44. A method according to claim 23, further comprising speaker verification means for comparing said determined parameter values with pre-stored speaker models to generate a verification result.

10       45. A speech processing apparatus comprising:  
         means for receiving a set of signal values representative of a speech signal generated by a speech source as modified by a transmission channel between the speech source and the receiving means;

15       a memory for storing data defining a predetermined function derived from a predetermined signal model which includes a first part having first parameters which models said source and a second part having second parameters which models said channel, said function being  
20       in terms of said first and second parameters and generating, for a given set of signal values, a probability density function which defines, for a given set of first parameters and second parameters, the probability that the predetermined signal model has those  
25       parameter values, given that the signal model is assumed

to have generated the received set of signal values;

means for applying said set of received signal values to said function;

processing said function with those values applied to derive samples of at least said first parameters from said probability density function; and

means for analysing at least some of said derived samples to determine values of said first parameters that are representative of said speech signal generated by said source before it was modified by said transmission channel.

46. A speech processing apparatus comprising:

means for receiving a set of signal values representative of a speech signal generated by a speech source as distorted by a transmission channel between the speech source and the receiving means;

a memory for storing data defining a predetermined function derived from a predetermined signal model which includes a first part having first parameters which models said source and a second part having second parameters which models said channel, said function being in terms of said first and second parameters and being in terms of a set of raw speech signal values representative of speech generated by said source before being distorted

by said transmission channel;

means for processing said received set of signal values with initial estimates of said first and second parameters, to generate an estimate of the raw speech signal values corresponding to the received set of signal values;

means for applying said set of received signal values and the estimated set of raw speech signal values to said function;

means for processing said function with those values applied to obtain values of said first parameters that are representative of said speech signal generated by said speech source before it was distorted by said transmission channel.

47. A speech processing apparatus comprising:

means for receiving a set of signal values representative of a speech signal generated by a speech source as modified by a transmission channel between the signal source and the receiving means;

a memory for storing data defining a predetermined function derived from a predetermined signal model which includes a first part having first parameters which models said source and a second part having second parameters which models said channel, said function being in terms

of said first and second parameters and being in terms of a set of raw speech signal values representative of a speech signal generated by said source before being modified by said transmission channel, said function generating, for a given set of signal values, a probability density function which defines, for a given set of first parameters, second parameters and raw speech signal values, the probability that the predetermined signal model has those parameter values and generates said raw speech signal values, given that the signal model is assumed to have generated the received set of signal values;

means for processing said received set of signal values with initial estimates of said first and second parameters, to generate an estimate of said set of raw speech signal values corresponding to the received set of signal values;

means for applying said set of received signal values and the estimated set of raw speech signal values to said function;

means for processing said function with those values applied to derive samples of at least said first parameters from said probability density function; and

means for analysing at least some of said derived samples to determine values of said first parameters that



are representative of said speech signal generated by said speech source before it was modified by said transmission channel.

5 48. An apparatus for determining sets of parameter values representative of an input speech signal, the apparatus comprising:

means for receiving a plurality of speech signal values representative of an input speech signal;

10 means for dividing the plurality of speech signal values into a succession of groups of speech signal values; and

means for determining a set of parameter values representative of the speech signal values in each group;

15 wherein said determining means comprises means for varying the number of parameter values used to represent the speech signal values in each group.

20 49. An apparatus for determining sets of parameter values representative of an input speech signal, the apparatus comprising:

means for receiving a plurality of speech signal values representative of an input speech signal;

25 means for dividing the plurality of speech signal values into a succession of groups of speech signal

values; and

means for processing the speech signal values in each group to determine a set of parameter values representative of the speech signal values in the group;

5        wherein said processing means comprises:

10        a memory for storing data defining a predetermined function which gives, for a set of speech signal values of a group, a probability density for parameters of a predetermined signal model which is assumed to have generated the speech signal values in the group, the probability density defining, for a given set of parameter values, the probability that the predetermined signal model has those parameter values, given that the model is assumed to have generated the speech signal values in the group;

15

means for applying the set of speech signal values of a current group to said stored function to give the probability density for said model parameters for the current group;

20        means for processing said function to derive samples of parameter values from said probability density for the current group;

25        means for evaluating said probability density for the current group using one or more of said derived samples of parameter values for different numbers of parameter

values to determine respective probabilities that the predetermined signal model has those parameter values; and

means for processing at least some of said derived samples of parameter values and said evaluated probabilities to determine model parameters that are representative of the set of signal values in the current group.

50. An audio comparison apparatus comprising:

a memory for storing a predetermined function which gives, for a given set of audio signal values, a probability density for parameters of a predetermined audio model which is assumed to have generated the set of audio signal values, the probability density defining, for a given set of model parameter values, the probability that the predetermined audio model has those parameter values, given that the model is assumed to have generated the set of audio signal values;

means for receiving a set of audio signal values representative of an input audio signal;

means for applying the set of received audio signal values to said stored function to give the probability density for said model parameters for the set of received audio signal values;

means for processing said function, with said set of

received audio signal values applied, to derive samples of parameter values from said probability density;

means for analysing at least some of said derived samples of parameter values to determine parameter values that are representative of the set of received audio signal values; and

means for comparing said determined parameter values with pre-stored parameter values to generate a comparison result.

51. A speech recognition system comprising an audio comparison apparatus according to claim 50.

52. A speaker verification system comprising an audio comparison apparatus according to claim 50.

53. An acoustic classification system comprising an audio comparison apparatus according to claim 50.

54. A computer readable medium storing computer executable process steps to cause a programmable computer apparatus to perform the method of claim 23.

55. Processor implementable process steps for causing a programmable computing device to perform the method according to claim 23.